ehp ENVIRONMENTAL HEALTH PERSPECTIVES
http://www.ehponline.org

Association of Cardiometabolic Genes with Arsenic Metabolism Biomarkers in American Indian Communities: The Strong Heart Family Study (SHFS)

Poojitha Balakrishnan, Dhananjay Vaidya, Nora Franceschini, V. Saroja Voruganti, Matthew O. Gribble, Karin Haack, Sandra Laston, Jason G. Umans, Kevin A. Francesconi, Walter Goessler, Kari E. North, Elisa Lee, Joseph Yracheta, Lyle G. Best, Jean W. MacCluer, Jack Kent, Jr, Shelley A. Cole, and Ana Navas-Acien

http://dx.doi.org/10.1289/EHP251

Received: 2 November 2015
Revised: 27 March 2016
Accepted: 19 May 2016
Published: 28 June 2016

Environ Health Perspect DOI: 10.1289/EHP251
Advance Publication: Not Copyedited

Association of Cardiometabolic Genes with Arsenic Metabolism Biomarkers in American Indian Communities: The Strong Heart Family Study (SHFS)

Poojitha Balakrishnan[1-3], Dhananjay Vaidya[2,4], Nora Franceschini[5], V. Saroja Voruganti[6], Matthew O. Gribble[7], Karin Haack[8], Sandra Laston[9], Jason G. Umans[10,11], Kevin A. Francesconi[12], Walter Goessler[12], Kari E. North[5], Elisa Lee[13], Joseph Yracheta[14], Lyle G. Best[14], Jean W. MacCluer[8], Jack Kent, Jr[8], Shelley A. Cole[8], and Ana Navas-Acien[1-3,15]

[1] Department of Environmental Health Sciences, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland;

[2] Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland;

[3] The Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins Medical Institutions, Baltimore, Maryland;

[4] Clinical and Translational Research, Johns Hopkins School of Medicine, Baltimore, Maryland;

[5] Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina;

[6] Department of Nutrition and UNC Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, North Carolina;

[7] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia;

[8] Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas;

[9] South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley, Brownsville, Texas;

[10] MedStar Health Research Institute, Hyattsville, Maryland;

[11] Georgetown and Howard Universities Center for Clinical and Translational Science, Washington, DC;

[12] Institute of Chemistry-Analytical Chemistry, University of Graz, Austria.

[13] Center for American Indian Health Research, College of Public Health, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma;

[14] Missouri Breaks Industries Research, Inc., Timber Lake, South Dakota;

[15] Department of Oncology, Johns Hopkins School of Medicine, Baltimore, Maryland;

Corresponding author:

Poojitha Balakrishnan

615 North Wolfe Street E7035

Baltimore, MD 21202

United States

Phone: 510-417-0783

Email: pbalakr2@jhu.edu

Short running title: Genetics of Arsenic Metabolism Biomarkers

Acknowledgement: The authors would like to thank the participants and staff of the Strong Heart and Strong Heart Family Study for their important contributions. The opinions expressed in this paper are those of the author(s) and do not necessarily reflect the view of the IHS. Work funded by National Institute of Environmental Health Sciences (R01ES021367, T32 ES007141-32). JY is employed by Missouri Breaks Industries Inc, Timber Lake, South Dakota.

Competing financial interests: The authors declare they have no actual or potential competing financial interests.

Abstract

Background: Metabolism of inorganic arsenic (iAs) is subject to inter-individual variability, explained party by genetic determinants.

Objectives: To investigate the association of genetic variants with arsenic species and principle components of arsenic species in the Strong Heart Family Study (SHFS).

Methods: We examined variants previously associated with cardiometabolic traits (~200,000 from Illumina Cardio MetaboChip) or arsenic metabolism and toxicity (670) among 2,428 American Indian participants in SHFS. Urine arsenic species were measured by HPLC-ICPMS, and percent arsenic species (iAs, monomethylarsonate [MMA], and dimethylarsinate [DMA], divided by their sum x 100) were logit transformed. We created two orthogonal principal components that summarized iAs, MMA, and DMA and were also phenotypes for genetic analyses. Linear regression was performed for each phenotype, dependent on allele dosage of the variant. Models accounted for familial relatedness and were adjusted for age, sex, total arsenic levels, and population stratification. SNP associations were stratified by study site and were meta-analyzed. Bonferroni correction was used to account for multiple testing.

Results: Variants at 10q24 were statistically significant for all percent arsenic species and principal components of arsenic species. The index SNP for iAs%, MMA%, and DMA% (rs12768205) and for the principal components (rs3740394, rs3740393) were located near *AS3MT*, whose gene product catalyzes methylation of iAs to MMA and DMA. Among the candidate arsenic variant associations, functional SNPs in *AS3MT* and 10q24 were most significant ($P < 9.33e-5$).

Conclusions: This hypothesis-driven association study supports the role of common variants in arsenic metabolism, particularly *AS3MT* and 10q24.

Introduction

Inorganic arsenic (iAs) is a toxic and carcinogenic metalloid, found in groundwater, soil, food and air (Welch, Lico, Hughes 1988). Experimental and epidemiological studies support the role of chronic iAs exposure to arsenic toxicity (Hughes 2002). In particular, iAs exposure has been associated with cancer (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, World Health Organization, International Agency for Research on Cancer 2004), cardiometabolic disease (Moon, Guallar, Navas-Acien 2012; Navas-Acien et al. 2008), and kidney disease (Peters et al. 2014; Zheng et al. 2015). After entering the body, iAs undergoes two sets of reduction and methylation to trivalent state and oxidation to a pentavalent state, producing monomethylarsonate (MMA) and then dimethylarsinate (DMA) (Hughes 2002). The arsenic species iAs, MMA, and DMA excreted via urine serve as biomarkers of arsenic metabolism (Vahter 1999). Although the pathways of arsenic toxicity are not completely established, trivalent arsenic metabolites have been implicated in arsenic related toxicity via mechanisms including epigenetic regulation, cytotoxicity, interfering with DNA repair, and oxidative stress (Hughes 2002).

More research is needed to understand the mechanisms of arsenic toxicity, especially given its systemic physiological effect on multiple organs. Arsenic toxicity may be better understood by investigating differences in response to arsenic; arsenic metabolism and patterns of arsenic methylation are subject to inter-individual variation, possibly influenced by genetic susceptibility, age, sex, nutrition, route of exposure, and other risk factors (Concha et al. 2002; Vahter 1999). The proportion of variation explained by genetic determinants, heritability, range from 50-53% for iAs%, 16-50% for MMA%, and 33-63% for DMA% (Chung et al. 2002; Gao et

al. 2015). Association studies have shown involvement of single nucleotide polymorphisms (SNPs) in arsenic (III) methyltransferase gene *AS3MT*, which encodes a major enzyme in biotransformation of iAs to MMA and DMA (Agusa et al. 2009; Rodrigues et al. 2012; Schläwicke Engström et al. 2009). More recently, genome-wide association studies (GWAS) in Bangladesh showed the association of *AS3MT* variants with MMA% and DMA% (Pierce et al. 2012). Additionally, since kidney disease, type 2 diabetes and other cardiometabolic diseases have been associated with arsenic exposure, genes associated with metabolic traits may be relevant determinants of arsenic species.

US Environmental Protection Agency (EPA) regulations state a maximum arsenic contaminant level in drinking water of 10 μg/L. In 2001, the US EPA Regional Tribal Program estimated 16.5% of tribally owned community water systems had at least one source with arsenic levels >10 μg/L, similar to other rural and suburban communities around the US, especially in the West, Midwest and Northeast (Navas-Acien et al. 2009). In this study, we characterized the genetic architecture of arsenic metabolism in extended American Indian families from Arizona, Oklahoma and North and South Dakota. Our objective was to replicate and to find novel associations of common variants with markers of arsenic metabolism, which can highlight biological pathways that influence arsenic toxicity and related diseases. Also as the evaluation of arsenic metabolism is complex being based on three interrelated biomarkers that sum to 100% (iAs%, MMA%, DMA%), we used a relatively novel method to summarize inter-individual variability in urine arsenic species patterns that are orthogonal using principal components analysis (PCA).

Methods

Study population

The Strong Heart Family Study (SHFS) is a large, multigenerational cohort recruited from the Strong Heart Study (SHS), an ongoing population-based study conducted in 13 American Indian tribes/communities in Arizona, Oklahoma and North/South Dakota. Details of the study design have been described previously (Lee et al. 1990; North et al. 2002). Briefly, families were eligible if they had a core sibship consisting of 3 original SHS participants and at least 5 additional living family members including 3 original SHS participants. Baseline visit was conducted in 2 phases; around 1,000 participants had their baseline visit in 1998-99 and over 2,500 participants had their baseline visit in 2001-03. Our study was restricted to 2,428 participants who were free of diabetes at the baseline visit, had urine arsenic species measured at baseline, were genotyped (MetaboChip and custom panel), and passed genotyping quality control, which is described in detail below. The SHFS protocols were approved by the Indian Health Service Institutional Review Board, by the Institutional Review Boards of the participating Institutions, and by the participating Indian tribes. Informed consent was obtained from all participants.

Arsenic measurements

Total urine arsenic and arsenic species concentrations were measured in spot urine samples collected the morning of the baseline visit. The samples were frozen and stored at -70^{o}C. Total urine arsenic was determined by inductively coupled plasma mass spectrometry (ICPMS) and arsenic species concentrations were determined by high performance liquid chromatography-ICPMS (HPLC-ICPMS) (Scheer et al. 2012). The inter-batch variability was monitored by

replicate measurements of three urine reference materials with certified arsenic levels of 20.3 µg/L (NIST 2669 I), 50.2 µg/L (NIST 2669 II) or 119 µg/L (NIES 18); the coefficients of variation ranged from 3.8 to 14.4%, with an overall mean cv of 7.9% (n=46) (**Supplemental Material Table S1**). The limit of quantitation for the total arsenic and arsenic species was 0.10 µg/L. There were 221 samples below the limit of detection for iAs (9.1%), 63 samples for MMA (2.6%), and 1 sample for DMA (<0.1%). The arsenic species levels below limit of quantitation were imputed as the limit of quantitation divided by the square root of 2 (0.07 µg/L). Since the lab assay included oxidization, the pentavalent and trivalent species were indistinguishable. Percent of each arsenic species (iAs%, MMA%, DMA%) was calculated as the relative proportion of the species to the sum of all three arsenic species.

SNP genotyping

DNA was extracted from blood specimens from the baseline visit using organic solvents (North et al. 2003) and was genotyped using the Illumina Cardio-Metabo DNA Analysis BeadChip (MetaboChip) containing 196,725 markers according to Illumina protocol (Voight et al. 2012). These markers were selected based on large-scale meta-analysis for cardiometabolic traits such as coronary artery disease and type 2 diabetes. Around one third of the MetaboChip SNPs consist of replication targets and almost two thirds are located in fine mapping regions, including AS3MT and other genes in the 10q24 region. Prior to genotyping quality control, non-autosomal and monomorphic markers were removed. Genotyping inconsistencies (Mendelian errors), were removed using Preswalk, a PEDSYS-compatible (Dyke 1996) version of Simwalk2 (Sobel, Papp, Lange 2002)and allele frequency estimation and Hardy Weinberg equilibrium (HWE) were estimated using Sequential Oligogenic Linkage Analysis Routines (SOLAR)

(Blangero and Almasy 1996) (Blangero and Almasy 1996). Family-based imputation was done using a PEDSYS-compatible version of Merlin (Abecasis et al. 2002). Participants were excluded if genotyping call rate was <95% (n=3). The SNP exclusion criteria included a call rate <98% or no data (n=33,604), not autosomal (n=250), monomorphic (n=158), HWE P <1e-5 (n=1,519), and minor allele frequency (MAF) <0.01 (n=40,219). As a result, there were 120,975 common variants used in the analysis. Pairwise correlations (r^2) between markers were calculated to estimate linkage disequilibrium (LD). A custom panel was used to genotype loci that were associated with arsenic traits in previous studies that were not already genotyped on the MetaboChip. A total of 670 arsenic candidate SNPs from 55 candidate genes (**Supplemental Material Table S2**) were genotyped. SNPs were assessed for assay inconsistencies and whether they were monomorphic in the sample. Samples were also assessed for genotyping errors using a call rate <95%, mismatch between genotyped and reported sex, outlier in identity by descent (IBD) clustering, or outlier in PCA. There were no SNPs that failed quality control.

Statistical Analysis

Percent arsenic species were logit transformed to approximate a normal distribution. Since percent arsenic species are interdependent, we also used PCA to summarize orthogonal dimensions of inter-individual variability in urine arsenic species patterns using the covariance structure of the arsenic species. Association analyses of 2,428 participants using MetaboChip SNPs and arsenic candidate SNPs were performed. All traits were modeled using linear regression of allele dosage at each SNP and adjusted for age at baseline, sex, total arsenic levels, and principal components for population stratification. An additive SNP effect was assumed. All analyses were stratified by study region and accounted for familial relatedness using SOLAR.

Inverse-variance-weighted meta-analysis of the stratified associations was performed using METAL (Willer, Li, Abecasis 2010). The MetaboChip-wide significance threshold was adjusted for multiple testing using Bonferroni correction (0.05/120,975 = 4.13e-7). Since the Bonferroni method would overcorrect for multiple testing in the presence of LD (Nyholt 2004), the suggestive threshold was calculated for the effective number of SNPs accounting for LD using SOLAR (0.05/64374.85 = 7.77e-7). Similarly for the arsenic candidate SNPs, the Bonferroni-corrected significance threshold was 9.33e-5 and the LD-corrected significance threshold was 9.1e-5. Association analysis conditioned on index SNP, the most statistically significant, informative SNP at the locus, was also performed using SOLAR. Simple linear regression analysis and stratification by sex were also performed for the arsenic traits as secondary analyses. All descriptive analysis was done using R version 3.2.2 (R Foundation for Statistical Computing).

Results

The median (IQR) for the sum of inorganic and methylated arsenic species was 6.6 (3.9-11.6) μg/L. Urine arsenic concentrations were higher in participants from Arizona compared to Oklahoma and North/South Dakota (**Table 1**). For arsenic metabolism, the median (IQR) was 9.8 (6.4-14.0) for iAs%, 13.9 (10.5-17.7) for MMA%, and 75.6 (68.6-81.6) for DMA%, with some variability across study regions (highest iAs% in Arizona, highest MMA% in North/South Dakota and highest DMA% in Oklahoma). The variability in iAs%, MMA% and DMA% can be summarized in two principal components. Principal component 1 (PC1) explained 86.1% of the variance in arsenic species and reflected higher iAs% and MMA% and lower DMA% (**Table 2**).

PC2 explained the remaining 13.9% of variance in arsenic species and reflected higher iAs% and lower MMA% independent of DMA% (**Table 2**).

The 10q24 region was statistically significantly associated with all logit-transformed percent arsenic species using the MetaboChip SNPs (**Figure 1**). The index SNP rs12768205 (G>A) in *AS3MT* was consistently associated with percent arsenic species (positively with iAs% and negatively with MMA%, DMA%) and principal components (negatively with PC1 and positively with PC2) (**Table 3, Supplemental Material Table S3-S7**). For PC1, SNPs in 10q24 passed the MetaboChip-wide alpha threshold and the index SNP rs3740394 (A>G) in *AS3MT* was also associated with DMA% (**Table 3**). An intronic SNP rs3740393 (C>G) in *AS3MT* was the top SNP for PC2 and was also significantly associated with MMA%, DMA% and PC1 (**Table 3**). There were other SNPs within the LD block with pairwise correlation r^2 >0.80 of the index SNP (104.62-104.65mb) that were statistically significantly associated with PC1 and PC2 (**Figure 2, Supplemental Material Table S6-S7**). Quantile-quantile plots (**Supplemental Material Figure S1**), Manhattan plots (**Supplemental Material Figure S3-S7**) and top SNP associations (**Supplemental Material Table S2-S6**) for MetaboChip SNPs are presented in the supplemental material. Association analyses in the 10q24 region conditioned on the index SNP did not yield any statistically significant independent associations that were not in LD (**Supplemental Material Figure S8-S9**).

Similar to the MetaboChip SNP associations, the only region that passed the multiple testing correction among the arsenic candidate SNPs was 10q24 (**Table 4, Supplemental Material Table S8-S12**). A total of 9 SNPs were associated with at least one trait; 5 SNPs were intronic and 4 SNPs had known functional changes. There were 2 coding SNPs (rs11191439,

T>C in *AS3MT*; rs4925, G>T in *GSTO1*) that alter the amino acid sequence and 2 SNPs (rs7911488, A>G in *USMG5*; rs2297235, A>G in *GSTO2*) that were located in the 3' untranslated region in the mRNA after a stop codon. The quantile-quantile plots (**Supplemental Material Figure S2**) and top SNP associations (**Supplemental Material Table S8-S12**) for the arsenic candidate SNPs are presented in the supplemental material. Secondary analyses including the unadjusted models, minimally adjusted models, stratified models by sex, and the conditional models for the arsenic traits did not yield any additional significant SNP associations.

The index SNP rs12768205 explained 1-15% of the heritability of the trait: iAs% (Arizona 5.8%, Oklahoma <0.1%, North/South Dakota 1.1%), MMA% (Arizona 10.1%, Oklahoma 0.5%, North/South Dakota 2.2%), DMA% (Arizona 12.1%, Oklahoma 14.8%, North/South Dakota 3.4%). At rs12768205, each copy of the variant allele G had a separation in the distribution of percent arsenic species (**Figure 3**). The variant genotype GG had iAs% and MMA% distributions that were shifted towards lower percentages and DMA% distribution that was shifter towards higher percentages. There was also a clear additive effect with each copy of the variant allele G for each percent arsenic species. The pattern of percent arsenic species by rs12768205 was similar when stratified by study region. The distribution of percent arsenic species by the index SNP genotype for PC1 (**Supplemental Material Figure S10**) and PC2 (**Supplemental Material Figure S11**) are presented in the supplemental material.

Discussion

We examined common variants (MAF >1%) in arsenic metabolism among American Indians as measured by logit-transformed iAs%, MMA%, DMA%, and principal components of logit arsenic species (PC1, PC2). The variants evaluated were available on the MetaboChip with

almost 200,000 markers supplemented with a custom panel of 670 candidate SNPs for arsenic metabolism and toxicity. Locus 10q24 was consistently associated with the arsenic phenotypes, similar to previous GWAS among unrelated individuals in Bangladesh and Vietnam (Agusa et al. 2009; Pierce et al. 2012; Rodrigues et al. 2012) and with a linkage peak in chromosome 10 among SHFS participants (Gribble et al. 2015). In our study, *AS3MT* SNPs produced significant estimates in 10q24 locus and the results suggest that common variants play a substantial role in the variation of percent arsenic species and principal components of arsenic species.

In *AS3MT*, rs12768205 was associated with iAs% (P 8.27e-8), MMA% (P 1.20e-15), and DMA% (P 5.90e-24). The index SNPs for the principal components were also located within *AS3MT*. *AS3MT* encodes enzyme arsenic (III) methyltransferase, responsible for methylation of iAs to MMA/DMA. Prior studies show association with SNPs in LD with rs12768205 and arsenic species and skin lesions, clinical manifestation specific to high-chronic arsenic toxicity (Gao et al. 2015; Pierce et al. 2012; Rodrigues et al. 2012). Recently, a population genetics study indicated a positive selection force for *AS3MT* variants in Argentina exposed to high levels of arsenic in drinking water and characterized by higher urine DMA% (Schlebusch et al. 2015). It is unknown if this positive selection is specific to this community or present in other populations chronically exposed to arsenic. Our association analysis results indicate the same loci and possibly genes including *AS3MT*, *CNNM2*, and *GSTO1* may be involved in arsenic metabolism. The distribution of percent arsenic species showed differentiation by rs12768205 genotype, suggesting that a functional SNP in LD may be a major causal factor in arsenic metabolism. Interestingly, previous GWAS of blood pressure and schizophrenia have found associations with *AS3MT* variants and 10q24 locus (Cross-Disorder Group of the Psychiatric Genomics

Consortium 2013; Newton-Cheh et al. 2009), indicating a possible mechanism of arsenic metabolism and toxicity. It is also possible that AS3MT is involved in other metabolic processes not yet identified beyond arsenic metabolism.

Among the candidate arsenic SNPs, functional and potentially causal variants in *AS3MT* were associated consistently with iAs%, MMA%, and DMA%. In particular, a potential candidate, the missense SNP rs11191439 (Methionine→Threonine) in *AS3MT,* is located within the LD block for rs12768205, the index SNP for percent arsenic species traits. Non-*AS3MT* genes showed signals in our study as well among the candidate arsenic SNPs. An upstream SNP at *USMG5* was associated with all arsenic traits. *USMG5* is a protein coding gene that is involved in skeletal muscle growth (Meyer et al. 2007; Zhao et al. 2014). The function of *USMG5* is unknown but may involve interactions with ATP synthases (Meyer et al. 2007). Interestingly, phosphorylation of proteins encoded by *USMG5* are known to be upregulated by insulin *in vivo* in human skeletal cells (Meyer et al. 2007). Also, mice exposed to 100 μg/L arsenite for 5 weeks showed impaired muscle function and mitochondrial myopathy compared to controls (Ambrosio et al. 2014). Confirmation is needed on the relevance of *USMG5* for arsenic metabolism, through fine mapping or sequencing as well as through functional analyses.

A coding variant (rs4925) at *GSTO1* and an upstream variant (rs2297235) at *GSTO2* were associated with PC1. The SNP rs4925 was also associated with PC2. The glutathione S-transferase omega family of genes are related to transferring thiol functional groups and therefore of importance in phase II metabolism of many xenobiotics. Due to its hypothesized role in oxidative stress and carcinogenesis, polymorphisms in the GSTO family have been candidates for various cancers and late onset Alzheimer's disease (Piacentini et al. 2012; Santos and Ward

2008). In a population study based in Bangladesh, *GSTO* variants have been previously associated with urinary arsenic species (Rodrigues et al. 2012). While the GSTO family of proteins is implicated in the arsenic reduction pathway (Hernandez et al. 2008), we cannot infer causality from our candidate gene association results due to the strong LD in the region. Additional mechanistic and epidemiological research is needed to confirm the relevance of GSTO variants in arsenic metabolism and toxicity. We also cannot discount the possibility that in addition to the strong LD, the genotype at one variant/gene can affect the expression at another. In fact, *AS3MT* variants may matter for gene expression in the SHS (Gribble et al. 2014). Additional mechanistic and epidemiological research is needed to confirm the relevance of GSTO and other functional variants in arsenic metabolism and toxicity.

The biotransformation pathway for inorganic arsenic is unresolved; prior biological evidence suggests that arsenic metabolism is mainly determined by reduction followed by oxidative methylation (Cullen and Reimer 1989; Vahter and Concha 2001; Vahter 1999). Our results for principal components support this hypothesis. Since iAs% and MMA% are inversely related to DMA%, PC1 may represent the overall methylation to DMA. Since PC2 had an inverse relationship between iAs% and MMA% independent of DMA%, PC2 may represent the first methylation step from iAs to MMA. Also, the development of arsenic related chronic diseases may be related to reactive oxygen species and reactive nitrogen species due to arsenic toxicity (Jomova et al. 2011; Jomova and Valko 2011; Shi, Shi, Liu 2004; Valko et al. 2007). Metabolizing inorganic arsenic leads to damaging effects to most organs (Shi, Shi, Liu 2004). An imbalance of reactive oxygen or nitrogen species exceeding the body's physiological antioxidant defenses can lead to widespread tissue injury, organ dysfunction, and clinical disease via

oxidative stress (Jomova et al. 2011; Valko et al. 2007). The genes associated with arsenic species and principal components of arsenic species in our analysis support this oxidative stress hypothesis (Jomova et al. 2011). In particular, prior studies demonstrate that AS3MT and non-AS3MT proteins such as USMG and GSTO can reduce pentavalent arsenic and facilitate transfer of arsenic intermediates as well as antioxidant depletion within and between cells (Hughes 2002; Lefort et al. 2009; Tanaka-Kagawa et al. 2003; Vahter and Concha 2001). These proteins may also result in oxidative stress via mitochondrial dysregulation from a build-up of free radicals (Lefort et al. 2009; Tanaka-Kagawa et al. 2003) and mitochondrial function being ubiquitous to all tissues may affect most organ systems and chronic diseases (Jomova et al. 2011; Valko et al. 2007).

This is the first large-scale study assessing markers of arsenic metabolism and common variants in a representative U.S. American Indian population sample and also one of the few studies evaluating genetic determinants of arsenic metabolism in a population exposed to low to moderate levels of arsenic. The Strong Heart Family Study presents a rich family-based cohort with low limits of quantification and limited missing data for arsenic exposure, arsenic species and other covariates. Another strength of this study is the consistency of findings within our study and to existing literature. In our study, the index SNP rs12768205 located in the *AS3MT* gene was associated with iAs%, MMA% and DMA% and SNPs in LD with rs12768205 were associated with PC1 and PC2. AS3MT has been previously reported to influence arsenic traits, including total arsenic levels (Argos et al. 2011; Pierce et al. 2012). Although individual urine arsenic metabolites are used in most cohort studies to assess arsenic metabolism, the pattern of arsenic metabolites in blood is different from urine and it is unknown if the genetic determinants

for urine arsenic metabolism correspond to the genetic determinants of arsenic metabolism as measured in blood (Goullé et al. 2005; Nixon and Moyer 1996; Tellez-Plaza et al. 2013; Vahter 1999; Valentine, Kang, Spivey 1979). In particular although blood arsenic levels would be a more proximal biomarker of arsenic metabolism, they tend to be present at much lower levels that are difficult to detect by conventional spectrophotometric methods (Kristiansen et al. 1997). Furthermore, the few epidemiologic studies using blood arsenic measurements, conducted in populations exposed to high arsenic levels in drinking water, have shown consistent associations compared to urine arsenic measurements (Hall et al. 2006; Valentine, Kang, Spivey 1979). No studies, however, have evaluated the association of genetic variants with blood arsenic species. Future research is needed to evaluate if the genetic variants associated with arsenic metabolism measured in urine are similar or different as those associated with arsenic metabolism measured in blood. Finally, PCA allowed us to not only reduce the number of dimensions but also to account for the interdependence of arsenic species so that PC1 and PC2 represent independent traits.

The analysis of MetaboChip SNPs allowed us to test hypotheses-driven variants since MetaboChip SNPs were chosen due to prior evidence of association with cardiometabolic diseases. In addition we assessed SNPs that were previously associated with arsenic species and arsenic related traits. This resulted in an investigation of SNPs with higher biological plausibility than regular genome-wide association studies. However this also limited the ability to investigate novel SNP associations. Another limitation is that the MetaboChip was a panel built using results from European American and African American populations (Voight et al. 2012). The SNP coverage and LD patterns may be different when extrapolated to other populations although the

use of MetaboChip has been characterized among populations with Asian or Mexican ancestry (Crawford et al. 2013). In addition, we were not able to test rare variants. Given the large statistical significance and high LD observed in the 10q24 locus, it is possible that rare functional variants may be causal. We were able to look at some low frequency variants (MAF 1-5%) among the candidate genes but genome sequencing data may be more useful in identifying putative causal SNPs. It is also possible that the multiple testing correction was too strict and therefore type II error may be present especially in regions of strong LD such as locus 10q24, limiting our ability to find variants with possibly weaker effects compared to the *AS3MT* variants. While the SHFS has a relatively small sample size, especially for genetic association studies, the strength of the associations support the importance of investigating genetic variants of arsenic metabolism within a prospective American Indian cohort. Finally given the low to moderate levels of arsenic exposure in Arizona, Oklahoma and the Dakotas, the generalizability of our findings need to be further assessed in other populations.

Conclusions

Association signals in *AS3MT* and surrounding genes in 10q24 are consistently associated with percent arsenic species and principal components of arsenic species. Furthermore, functionally annotated variants in 10q24 also show a strong relationship with the arsenic traits. The associated genes such as *AS3MT* and *GSTO1/2* highlight oxidative stress as a possible mechanism in arsenic biotransformation and therefore arsenic related diseases. Due to the high LD in the 10q24 region in populations around the world (Fujihara et al. 2010; Gomez-Rubio et al. 2010) and especially among American Indians in our study, further investigation in comparable populations and using low frequency variants is needed to confirm our findings.

Further knowledge of causal variation may highlight biological mechanisms that are related to arsenic metabolism, including methylation, and contribute to understand possible mechanisms for arsenic toxicity and development of chronic diseases including skin lesions, cancer and cardiovascular disease.

References

Agusa T, Iwata H, Fujihara J, Kunito T, Takeshita H, Minh TB, Trang PTK, Viet PH, Tanabe S. 2009. Genetic polymorphisms in< i> AS3MT and arsenic metabolism in residents of the red river delta, vietnam. Toxicol Appl Pharmacol 236(2):131-41.

Ambrosio F, Brown E, Stolz D, Ferrari R, Goodpaster B, Deasy B, Distefano G, Roperti A, Cheikhi A, Garciafigueroa Y. 2014. Arsenic induces sustained impairment of skeletal muscle and muscle progenitor cell ultrastructure and bioenergetics. Free Radical Biology and Medicine 74:64-73.

Argos M, Kalra T, Pierce BL, Chen Y, Parvez F, Islam T, Ahmed A, Hasan R, Hasan K, Sarwar G, et al. 2011. A prospective study of arsenic exposure from drinking water and incidence of skin lesions in bangladesh. Am J Epidemiol 174(2):185-94.

Blangero J and Almasy L. 1996. Solar: Sequential oligogenic linkage analysis routines. Population Genetics Laboratory Technical Report 6.

Chung JS, Kalman DA, Moore LE, Kosnett MJ, Arroyo AP, Beeris M, Mazumder DN, Hernandez AL, Smith AH. 2002. Family correlations of arsenic methylation patterns in children and parents exposed to high concentrations of arsenic in drinking water. Environ Health Perspect 110(7):729-33.

Concha G, Vogler G, Nermell B, Vahter M. 2002. Intra-individual variation in the metabolism of inorganic arsenic. Int Arch Occup Environ Health 75(8):576-80.

Crawford Dana C., Goodloe Robert, Brown-Gentry Kristin, Wilson Sarah, Roberson Jamie, Gillani Niloufar B., Ritchie Marylyn D., Dilks Holli H. and Bush William S. 2013. Characterization of the metabochip in diverse populations from the international HapMap project in the epidemiologic architecture for genes linked to environment (EAGLE) project. Pac symp biocomput.

Cross-Disorder Group of the Psychiatric Genomics Consortium. 2013. Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. The Lancet 381(9875):1371-9.

Cullen WR and Reimer KJ. 1989. Arsenic speciation in the environment. Chem Rev 89(4):713-64.

Fujihara J, Soejima M, Yasuda T, Koda Y, Agusa T, Kunito T, Tongu M, Yamada T, Takeshita H. 2010. Global analysis of genetic variation in human arsenic (3 oxidation state) methyltransferase (AS3MT). Toxicol Appl Pharmacol 243(3):292-9.

Gao J, Tong L, Argos M, Bryan MS, Ahmed A, Rakibuz-Zaman M, Kibriya MG, Jasmine F, Slavkovich V, Graziano JH, et al. 2015. The genetic architecture of arsenic metabolism efficiency:A SNP-based heritability study of bangladeshi adults. Environ Health Perspect 123(10):985-92.

Gomez-Rubio P, Meza-Montenegro MM, Cantu-Soto E, Klimecki WT. 2010. Genetic association between intronic variants in AS3MT and arsenic methylation efficiency is focused on

a large linkage disequilibrium cluster in chromosome 10. Journal of Applied Toxicology 30(3):260-70.

Goullé J, Mahieu L, Castermant J, Neveu N, Bonneau L, Lainé G, Bouige D, Lacroix C. 2005. Metal and metalloid multi-elementary ICP-MS validation in whole blood, plasma, urine and hair: Reference values. Forensic Sci Int 153(1):39-44.

Gribble MO, Tang W, Shang Y, Pollak J, Umans JG, Francesconi KA, Goessler W, Silbergeld EK, Guallar E, Cole SA. 2014. Differential methylation of the arsenic (III) methyltransferase promoter according to arsenic exposure. Arch Toxicol 88(2):275-82.

Gribble MO, Voruganti VS, Cole SA, Haack K, Balakrishnan P, Laston SL, Tellez-Plaza M, Francesconi KA, Goessler W, Umans JG, et al. 2015. Linkage analysis of urine arsenic species patterns in the strong heart family study. Toxicol Sci 148(1):89-100.

Hall M, Chen Y, Ahsan H, Slavkovich V, Van Geen A, Parvez F, Graziano J. 2006. Blood arsenic as a biomarker of arsenic exposure: Results from a prospective study. Toxicology 225(2):225-33.

Hernandez A, Xamena N, Sekaran C, Tokunaga H, Sampayo-Reyes A, Quinteros D, Creus A, Marcos R. 2008. High arsenic metabolic efficiency in AS3MT287Thr allele carriers. Pharmacogenet Genomics 18(4):349-55.

Hughes MF. 2002. Arsenic toxicity and potential mechanisms of action. Toxicol Lett 133(1):1-16.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, World Health Organization, International Agency for Research on Cancer. 2004. Some drinking-water disinfectants and contaminants, including arsenic. IARC.

Jomova K, Jenisova Z, Feszterova M, Baros S, Liska J, Hudecova D, Rhodes C, Valko M. 2011. Arsenic: Toxicity, oxidative stress and human disease. Journal of Applied Toxicology 31(2):95-107.

Jomova K and Valko M. 2011. Advances in metal-induced oxidative stress and human disease. Toxicology 283(2):65-87.

Kristiansen J, Christensen JM, Iversen BS, Sabbioni E. 1997. Toxic trace element reference levels in blood and urine: Influence of gender and lifestyle factors. Sci Total Environ 204(2):147-60.

Lee ET, Welty TK, Fabsitz R, Cowan LD, Le NA, Oopik AJ, Cucchiara AJ, Savage PJ, Howard BV. 1990. The strong heart study. A study of cardiovascular disease in american indians: Design and methods. Am J Epidemiol 132(6):1141-55.

Lefort N, Yi Z, Bowen B, Glancy B, De Filippis EA, Mapes R, Hwang H, Flynn CR, Willis WT, Civitarese A. 2009. Proteome profile of functional mitochondria from human skeletal muscle using one-dimensional gel electrophoresis and HPLC-ESI-MS/MS. Journal of Proteomics 72(6):1046-60.

Meyer B, Wittig I, Trifilieff E, Karas M, Schagger H. 2007. Identification of two proteins associated with mammalian ATP synthase. Mol Cell Proteomics 6(10):1690-9.

Moon K, Guallar E, Navas-Acien A. 2012. Arsenic exposure and cardiovascular disease: An updated systematic review. Curr Atheroscler Rep 14(6):542-55.

Navas-Acien A, Silbergeld EK, Pastor-Barriuso R, Guallar E. 2008. Arsenic exposure and prevalence of type 2 diabetes in US adults. JAMA 300(7):814-22.

Navas-Acien A, Umans JG, Howard BV, Goessler W, Francesconi KA, Crainiceanu CM, Silbergeld EK, Guallar E. 2009. Urine arsenic concentrations and species excretion patterns in american indian communities over a 10-year period: The strong heart study. Environ Health Perspect 117(9):1428-33.

Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS, Zhao JH, Heath SC, Eyheramendy S. 2009. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet 41(6):666-76.

Nixon DE and Moyer TP. 1996. Routine clinical determination of lead, arsenic, cadmium, and thallium in urine and whole blood by inductively coupled plasma mass spectrometry. Spectrochimica Acta Part B: Atomic Spectroscopy 51(1):13-25.

North KE, Howard BV, Welty TK, Best LG, Lee ET, Yeh JL, Fabsitz RR, Roman MJ, MacCluer JW. 2003. Genetic and environmental contributions to cardiovascular disease risk in american indians: The strong heart family study. Am J Epidemiol 157(4):303-14.

North KE, MacCluer JW, Devereux RB, Howard BV, Welty TK, Best LG, Lee ET, Fabsitz RR, Roman MJ, Strong Heart Family Study. 2002. Heritability of carotid artery structure and function: The strong heart family study. Arterioscler Thromb Vasc Biol 22(10):1698-703.

Nyholt DR. 2004. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. The American Journal of Human Genetics 74(4):765-9.

Peters BA, Hall MN, Liu X, Neugut YD, Pilsner JR, Levy D, Ilievski V, Slavkovich V, Islam T, Factor-Litvak P. 2014. Creatinine, arsenic metabolism, and renal function in an arsenic-exposed population in bangladesh. 9(12):e113760.

Piacentini S, Polimanti R, Squitti R, Mariani S, Migliore S, Vernieri F, Rossini PM, Manfellotto D, Fuciarelli M. 2012. GSTO1* E155del polymorphism associated with increased risk for late-onset alzheimer's disease: Association hypothesis for an uncommon genetic variant. Neurosci Lett 506(2):203-7.

Pierce BL, Kibriya MG, Tong L, Jasmine F, Argos M, Roy S, Paul-Brutus R, Rahaman R, Rakibuz-Zaman M, Parvez F. 2012. Genome-wide association study identifies chromosome 10q24. 32 variants associated with arsenic metabolism and toxicity phenotypes in bangladesh. Plos Genetics 8(2):e1002522.

Rodrigues EG, Kile M, Hoffman E, Quamruzzaman Q, Rahman M, Mahiuddin G, Hsueh Y, Christiani DC. 2012. GSTO and AS3MT genetic polymorphisms and differences in urinary arsenic concentrations among residents in bangladesh. Biomarkers 17(3):240-7.

Santos L and Ward L. 2008. Lack of association of GSTT1, GSTM1, GSTO1, GSTP1 and CYP1A1 polymorphisms for susceptibility and outcome in brazilian prostate cancer patients. Folia Biologica (Praha) 54:102-8.

Scheer J, Findenig S, Goessler W, Francesconi KA, Howard B, Umans JG, Pollak J, Tellez-Plaza M, Silbergeld EK, Guallar E. 2012. Arsenic species and selected metals in human urine: Validation of HPLC/ICPMS and ICPMS procedures for a long-term population-based epidemiological study. Analytical Methods 4(2):406-13.

Schläwicke Engström K, Nermell B, Concha G, Strömberg U, Vahter M, Broberg K. 2009. Arsenic metabolism is influenced by polymorphisms in genes involved in one-carbon metabolism and reduction reactions. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 667(1):4-14.

Schlebusch CM, Gattepaille LM, Engstrom K, Vahter M, Jakobsson M, Broberg K. 2015. Human adaptation to arsenic-rich environments. Mol Biol Evol 32(6):1544-55.

Shi H, Shi X, Liu KJ. 2004. Oxidative mechanism of arsenic toxicity and carcinogenesis. Mol Cell Biochem 255(1-2):67-78.

Sobel E, Papp JC, Lange K. 2002. Detection and integration of genotyping errors in statistical genetics. The American Journal of Human Genetics 70(2):496-508.

Tanaka-Kagawa T, Jinno H, Hasegawa T, Makino Y, Seko Y, Hanioka N, Ando M. 2003. Functional characterization of two variant human GSTO 1-1s (Ala140Asp and Thr217Asn). Biochem Biophys Res Commun 301(2):516-20.

Tellez-Plaza M, Gribble MO, Voruganti VS, Francesconi KA, Goessler W, Umans JG, Silbergeld EK, Guallar E, Franceschini N, North KE, et al. 2013. Heritability and preliminary genome-wide linkage analysis of arsenic metabolites in urine. Environ Health Perspect 121(3):345-51.

Vahter M and Concha G. 2001. Role of metabolism in arsenic toxicity. Pharmacol Toxicol 89(1):1-5.

Vahter M. 1999. Methylation of inorganic arsenic in different mammalian species and population groups. Sci Prog 82 (Pt 1)(Pt 1):69-88.

Valentine JL, Kang HK, Spivey G. 1979. Arsenic levels in human blood, urine, and hair in response to exposure via drinking water. Environ Res 20(1):24-32.

Valko M, Leibfritz D, Moncol J, Cronin MT, Mazur M, Telser J. 2007. Free radicals and antioxidants in normal physiological functions and human disease. Int J Biochem Cell Biol 39(1):44-84.

Voight BF, Kang HM, Ding J, Palmer CD, Sidore C, Chines PS, Burtt NP, Fuchsberger C, Li Y, Erdmann J. 2012. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet 8(8):e1002793.

Welch AH, Lico MS, Hughes JL. 1988. Arsenic in ground water of the western united states. Groundwater 26(3):333-47.

Willer CJ, Li Y, Abecasis GR. 2010. METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26(17):2190-1.

Zhao X, Bak S, Pedersen AJ, Jensen ON, Højlund K. 2014. Insulin increases phosphorylation of mitochondrial proteins in human skeletal muscle in vivo. Journal of Proteome Research 13(5):2359-69.

Zheng LY, Umans JG, Yeh F, Francesconi KA, Goessler W, Silbergeld EK, Bandeen-Roche K, Guallar E, Howard BV, Weaver VM. 2015. The association of urine arsenic with prevalent and incident chronic kidney disease. Epidemiology 26(4):601-12.

Table 1. Baseline characteristics of Strong Heart Family Study participants

	Total	**Arizona**	**Oklahoma**	**South/North Dakota**	**P-value**
No. of Participants	2428	703	819	906	
Mean age, yrs (SD)	35.2 (15.2)	30.9 (13.1)	38.8 (15.7)	35.3 (15.4)	<0.01
No. females (%)	1469 (60.5%)	438 (62.3%)	486 (59.34%)	545 (60.15%)	0.48
Arsenic levels					<0.01
Tertile 1 (0.21- 4.73µg/L)	804 (33.1%)	100 (14.2%)	357 (43.5%)	347 (38.3%)	
Tertile 2 (4.74-9.35µg/L)	799 (32.9%)	204 (29.0%)	293 (35.7%)	302 (33.3%)	
Tertile 3 (9.37- 176.6µg/L)	825 (33.9%)	399 (56.7%)	169 (20.6%)	257 (28.3%)	
Median iAs% (IQR)	9.78 (6.4-14.0)	10.8 (7.6-15.2)	8.2 (5.4-12.3)	10.1 (6.7-14.5)	<0.01
Median MMA% (IQR)	13.9 (10.5-17.7)	12.8 (10.0-16.2)	13.7 (10.3-17.8)	14.9 (11.5-18.5)	<0.01
Median DMA% (IQR)	75.6 (68.6-81.6)	75.2 (69.0-81.1)	77.2 (70.5-83.3)	74.3 (67.3-80.4)	<0.01
Mean PC1 (SD)	0.04 (12.21)	0.62 (12.05)	-2.20 (11.64)	1.62 (12.54)	<0.01
Mean PC2 (SD)	-0.01 (4.91)	1.59 (4.87)	-0.70 (4.53)	-0.63 (5.00)	<0.01

Abbreviations: DMA%, percent dimethylarsinate. iAs%, percent inorganic arsenic. IQR, interquartile range. MMA%, percent monomethylarsonate. SD, standard deviation. yrs, years.

Table 2. Summary of principal components of arsenic species

	PC1	**PC2**
Variance in arsenic species explained (%)	86.1	13.9
Standard deviation	12.21	4.91
Weight for iAs%	0.49	0.65
Weight for MMA%	0.32	-0.75
Weight for DMA%	-0.81	0.10

Abbreviations: DMA%, percent dimethylarsinate. iAs%, percent inorganic arsenic. MMA%, percent monomethylarsonate. PC, principal components.

Table 3. Index MetaboChip SNPs of percent arsenic species and principal components of arsenic species

SNP	Chr	Position‡	Allele	MAF	Gene	Location	Trait	P-value				
								iAs%	MMA%	DMA%	PC1	PC2
rs3740393	10	104626645	C/G	0.17	AS3MT	intron	PC2	8.63e-7	1.07e-13*	2.71e-23*	2.58e-34*	1.56e-8*
rs3740394	10	104624464	A/G	0.18	AS3MT	intron	PC1	1.27e-4	4.66e-7†	3.87e-20*	2.19e-38*	5.57e-6
rs12768205	10	104637839	G/A	0.27	AS3MT	intron	iAs%, MMA%, DMA%	8.27e-8*	1.20e-15*	5.90e-24*	1.15e-29*	1.78e-7*

Abbreviations: Chr, chromosome. DMA%, percent dimethylarsinate. iAs%, percent inorganic arsenic. MMA%, percent monomethylarsonate. MAF, minor allele frequency. PC, principal components. SNP, single nucleotide polymorphism.

* Significant SNP associations (P<4.13e-7).

† Suggestive SNP associations (P<7.77e-7).

‡ Base position according to human genome build 18.

Table 4. Top Candidate SNPs of percent arsenic species and principal components of arsenic species

SNP	Chr	Position†	Allele	MAF	Gene	Location	P-value				
							iAs%	MMA%	DMA%	PC1	PC2
rs11191439	10	104638723	T/C	0.18	AS3MT	coding	3.23e-4	2.60e-7*	2.89e-19*	1.12e-36*	7.54e-7*
rs3740394	10	104634474	A/G	0.19	AS3MT	intron	2.53e-6*	7.55e-7*	8.30e-20*	2.83e-38*	6.09e-8*
rs3740390	10	104638480	C/T	0.20	AS3MT	intron	1.76e-8*	9.24e-13*	8.63e-23*	6.50e-34*	0.99
rs11191453	10	104659852	T/C	0.20	AS3MT	intron	6.48e-7*	5.57e-12*	2.18e-21*	1.80e-32*	0.34
rs4919694	10	104698978	T/C	0.18	CNNM2	intron	8.10e-10*	4.74e-7*	1.67e-18*	9.67e-36*	3.95e-8*
rs7911488	10	105154089	A/G	0.25	USMG5	UTR	1.04e-6*	6.54e-8*	7.04e-18*	3.57e-31*	6.63e-6*
rs4925	10	106022789	C/A	0.12	GSTO1	coding	0.02	0.11	1.55e-3	4.60e-6*	9.18e-7*
rs1147611	10	106025258	G/T	0.19	GSTO1	intron	0.11	0.35	7.23e-3	8.45e-6*	0.27
rs2297235	10	106034491	A/G	0.12	GSTO2	UTR	0.08	0.20	4.58e-3	1.67e-5*	4.31e-3

Abbreviations: Chr, chromosome. DMA%, percent dimethylarsinate. iAs%, percent inorganic arsenic. MMA%, percent monomethylarsonate. MAF, minor allele frequency. PC, principal components. SNP, single nucleotide polymorphism.

* Significant SNP associations (P<9.33e-5).

† Base position according to human genome build 18.

Figure Legends

Figure 1. Regional association plot at 10q24 of percent arsenic species

Index SNP rs12768205 nearby associations according to human genome build 18 for percent inorganic arsenic (iAs%) in top panel, percent monomethylarsonate (MMA%) in middle panel, and percent dimethylarsinate (DMA%) in bottom panel. The solid red line is the MetaboChip-wide significance threshold at -log(4.13e-7) or 6.38. The dashed orange line is the suggestive MetaboChip-wide LD threshold at –log(7.77e-7).

Figure 2. Regional association plot at 10q24 of principal components of arsenic species

Index SNPs for principal components (top panel PC1 rs3740394, bottom panel PC2 rs3740393) and nearby associations according to human genome build 18. The solid red line is the MetaboChip-wide significance threshold at -log(4.13e-7) or 6.38. The dashed orange line is the suggestive MetaboChip-wide LD threshold at –log(7.77e-7).

Figure 3. Distribution of percent arsenic species by rs12768205 genotype

Index SNP rs12768205 for percent arsenic species principal components shows separation of distribution of percent inorganic arsenic (iAs%), percent monomethylarsonate (MMA%) and dimethylarsinate (DMA%) by genotype. Of the 2,428 participants, the distribution of genotypes is homozygous major AA (1245), heterozygous AG (982), and homozygous minor GG (201).

Figure 1.

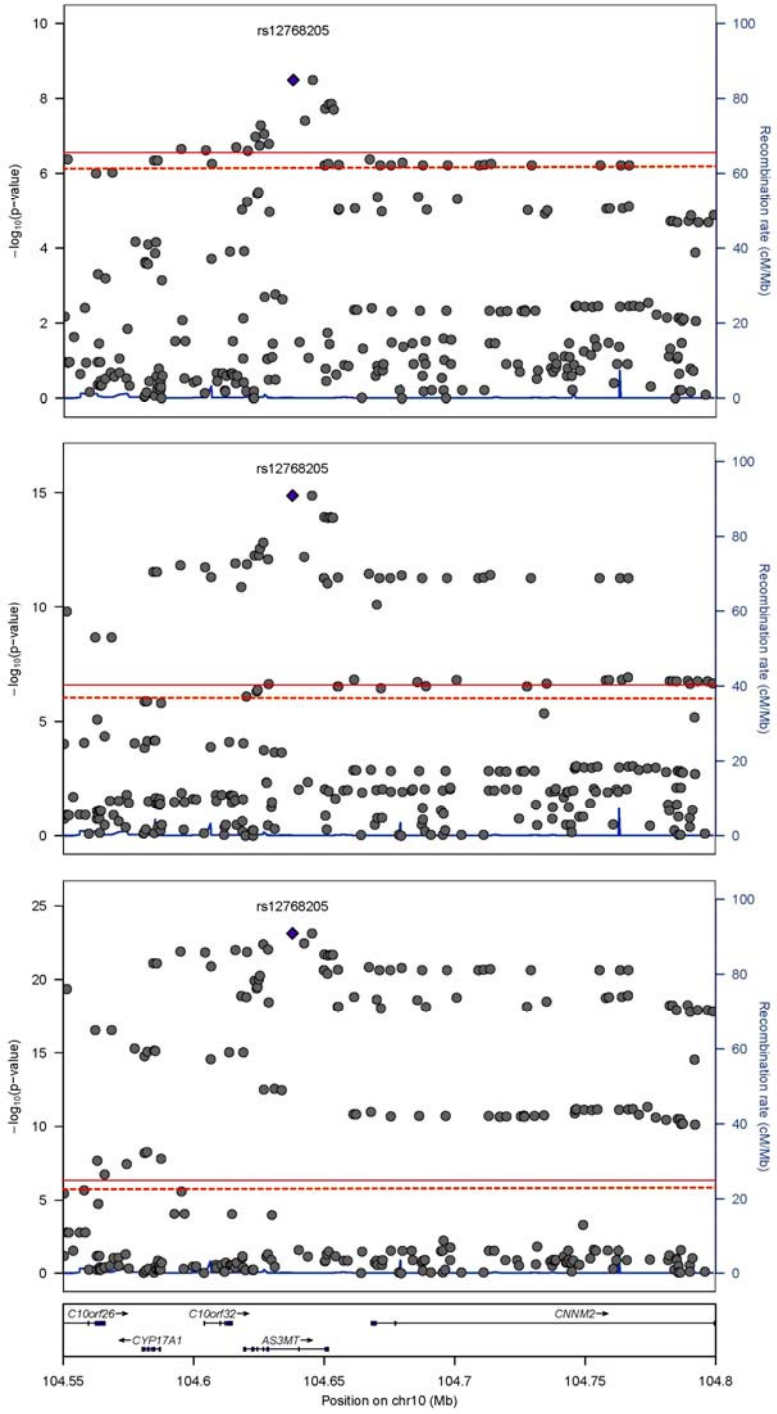

rs12768205
rs12768205
rs12768205
−log10(p−value)
Recombination rate (cM/Mb)
C10orf26
C10orf32
CNNM2
CYP17A1
AS3MT
104.55
104.6
104.65
104.7
104.75
104.8
Position on chr10 (Mb)

Figure 2.

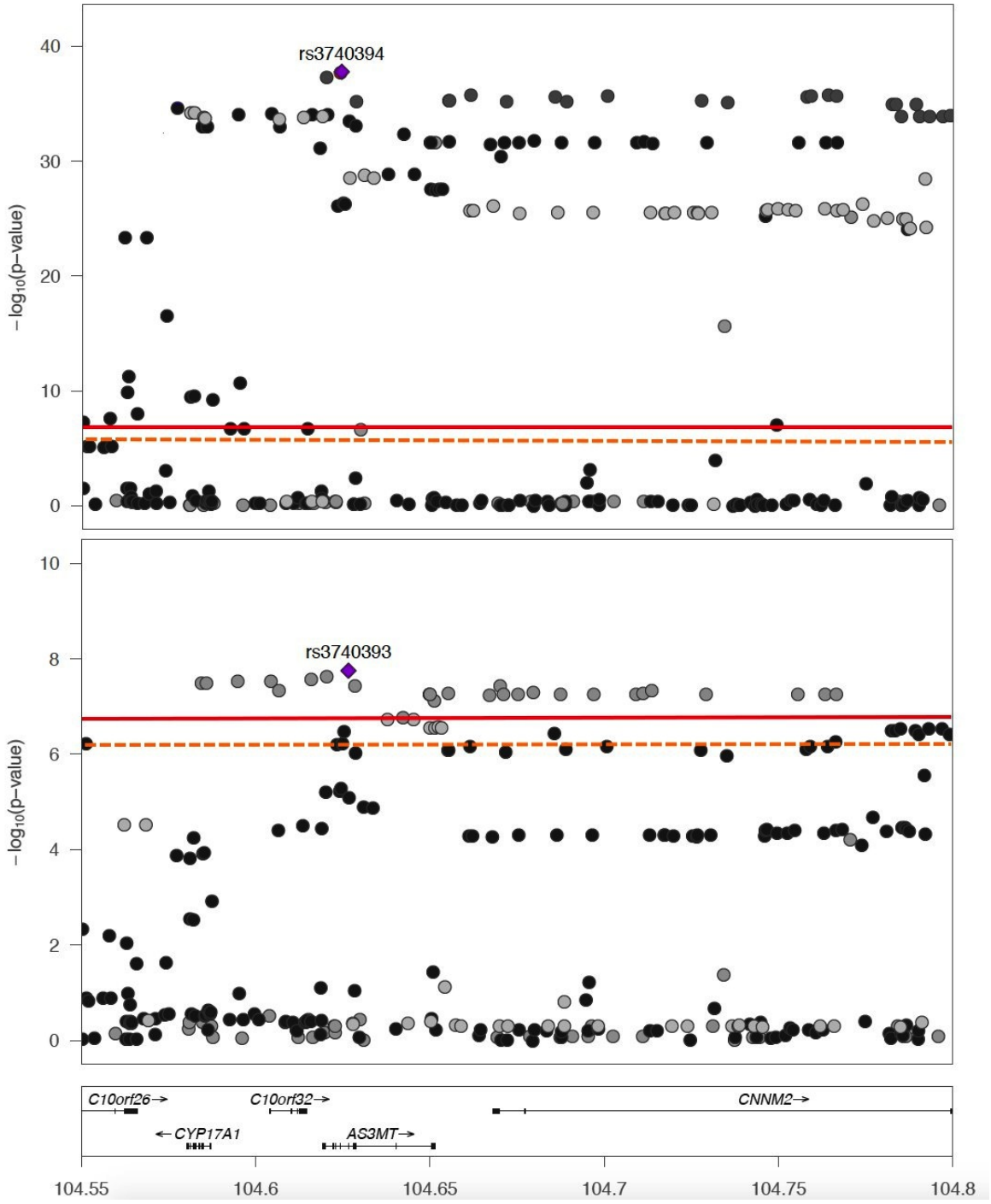
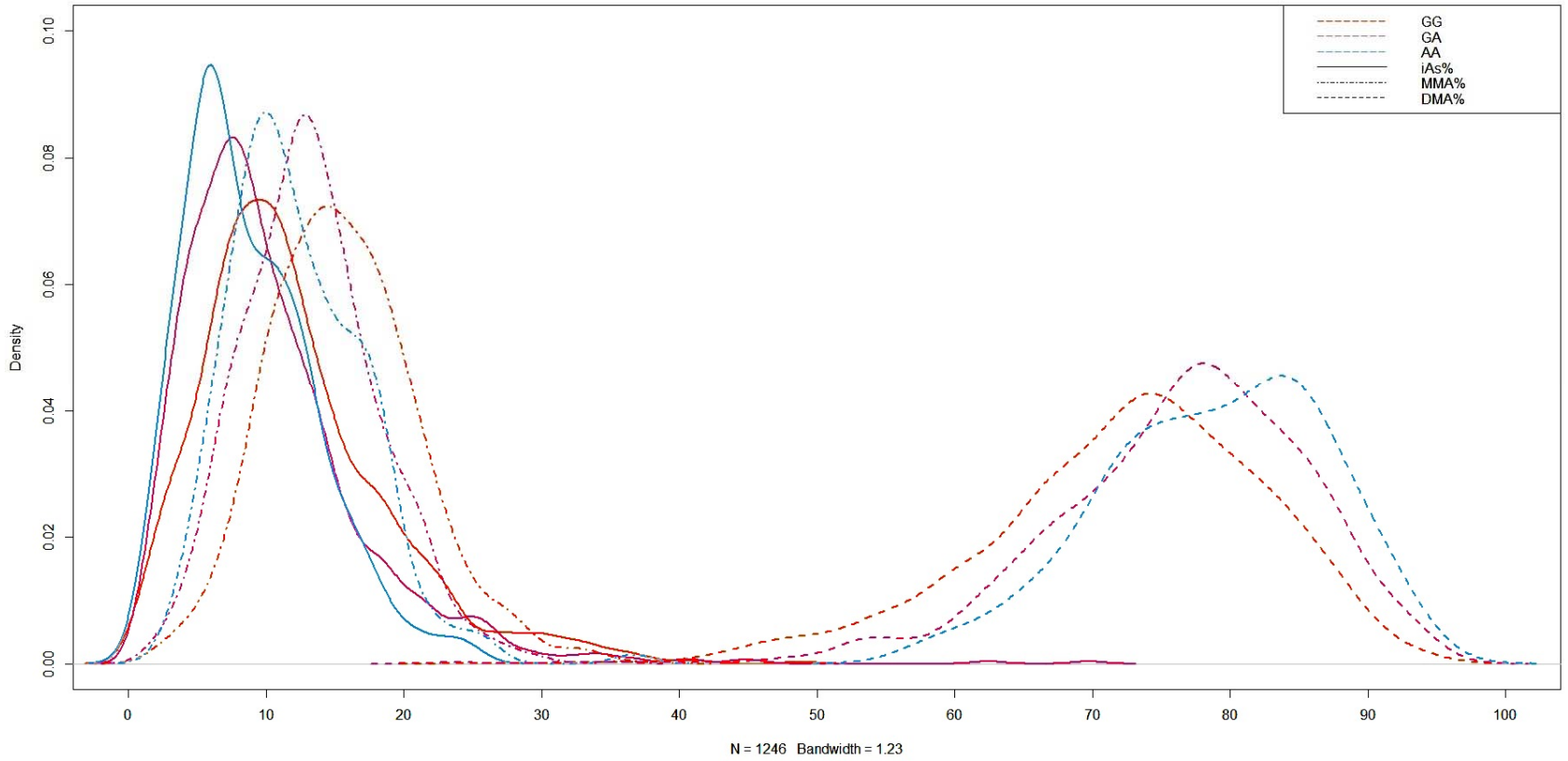

rs3740394
rs3740393
−log10(p−value)
C10orf26
C10orf32
CNNM2
CYP17A1
AS3MT
104.55
104.6
104.65
104.7
104.75
104.8

Figure 3.


GG
GA
AA
iAs%
MMA%
DMA%
Density
0.00
0.02
0.04
0.06
0.08
0.10
0
10
20
30
40
50
60
70
80
90
100
N = 1246 Bandwidth = 1.23